

## A STUDY OF SOME NINE-ELEMENT DECISION RULES

(NASA-CR-140203) A STUDY OF SOME  
NINE-ELEMENT DECISION RULES Final  
Report, 1 Feb. - 31 Oct. 1973

N74-31874

(Environmental Research Inst. of Michigan)  
46 p HC \$5.50

Unclas  
CSSL 05B G3/13 47883

by

Wyman Richardson  
Infrared and Optics Division



FORMERLY WILLOW RUN LABORATORIES,  
THE UNIVERSITY OF MICHIGAN

July 1974  
prepared for

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Johnson Space Center, Houston, Texas 77058  
Earth Observations Division  
Contract NAS 9-9784, Task VII



## NOTICES

Sponsorship. The work reported herein was conducted by the Environmental Research Institute of Michigan for the National Aeronautics and Space Administration, Johnson Space Center, Houston, Texas 77058 under contract NAS 9-9784, Task VII. Dr. Andrew Potter (TF3) is Technical Monitor. Contracts and grants to the Institute for the support of sponsored research are administered through the Office of Contracts Administration.

Disclaimers. This report was prepared as an account of Government-sponsored work. Neither the United States, nor the National Aeronautics and Space Administration (NASA), nor any person acting on behalf of NASA:

- (A) Makes any warranty or representation, expressed or implied with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or
- (B) Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used above, "person acting on behalf of NASA" includes any employee or contractor of NASA, or employee of such contractor, to the extent that such employee or contractor of NASA or employee of such contractor prepares, disseminates, or provides access to any information pursuant to his employment or contract with NASA, or his employment with such contractor.

Availability Notice. Requests for copies of this report should be referred to:

National Aeronautics and Space Administration  
Scientific and Technical Information Facility  
P. O. Box 33  
College Park, Maryland 20740

Final Disposition. After this document has served its purpose, it may be destroyed. Please do not return it to the Environmental Research Institute of Michigan.

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. <b>NASA CR-ERIM 190100-32-T</b>		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle <b>A STUDY OF SOME NINE-ELEMENT DECISION RULES</b>		5. Report Date <b>December 1973</b>			
		6. Performing Organization Code			
7. Author(s) <b>Wyman Richardson</b>		8. Performing Organization Report No. <b>ERIM 190100-32-T</b>			
9. Performing Organization Name and Address <b>Environmental Research Institute of Michigan Infrared and Optics Division P. O. Box 618 Ann Arbor, Michigan 48107</b>		10. Work Unit No. <b>Task VII</b>			
		11. Contract or Grant No. <b>NAS 9-9784</b>			
12. Sponsoring Agency Name and Address <b>National Aeronautics and Space Administration Johnson Space Center Earth Observations Division Houston, Texas 77058</b>		13. Type of Report and Period Covered <b>Final Report 1 February-31 October 1973</b>			
		14. Sponsoring Agency Code			
15. Supplementary Notes <b>Dr. Andrew Potter /TF-3 is Technical Monitor for NASA.</b>					
16. Abstract <p>A nine-element rule is one that makes a classification decision for each pixel based on data from that pixel and its eight immediate neighbors. Three such rules, all fast and simple to use, are defined and tested. All performed substantially better on field interiors than the best one-point rule. Qualitative results indicate that fine detail and contradictory testimony tend to be overlooked by the rules.</p>					
17. Key Words <b>Multi-element processing Remote sensing Classification rules Multispectral recognition Nine-element rules</b>			18. Distribution Statement <b>Initial distribution is listed at the end of this document.</b>		
19. Security Classif. (of this report) <b>UNCLASSIFIED</b>	20. Security Classif. (of this page) <b>UNCLASSIFIED</b>	21. No. of Pages <b>44</b>	22. Price		

## PREFACE

This report describes part of a comprehensive and continuing program of research concerned with advancing the state-of-the-art in remote sensing of the environment from aircraft and satellites. The research is being carried out for NASA's Lyndon B. Johnson Space Center, Houston, Texas, by the Environmental Research Institute of Michigan (ERIM), formerly the Willow Run Laboratories of The University of Michigan. The basic objective of this multidisciplinary program is to develop remote sensing as a practical tool to provide the planner and decision-maker with extensive information quickly and economically.

Timely information obtained by remote sensing can be important to such people as the farmer, the city planner, the conservationist, and others concerned with problems such as crop yield and disease, urban land studies and development, water pollution, and forest management. The scope of our program includes (1) extending the understanding of basic processes; (2) discovering new applications, developing advanced remote-sensing systems, and improving automatic data processing to extract information in a useful form; and (3) assisting in data collection, processing, analysis, and ground-truth verification.

The research described herein was performed under NASA Contract NAS 9-9784, Task VII and covers the period from February 1, 1973 through October 31, 1973. Dr. Andrew Potter has been Technical Monitor. The program was directed by R. R. Legault, Vice-President of ERIM, J. D. Erickson, Principal Investigator and Head of the Information Systems and Analysis Department, and R. F. Nalepka, Head of the Multispectral Analysis Section. The ERIM number for this report is 190100-32-T

The results reported in Appendix B were derived by H. M. Horwitz. R. B. Crane and R. J. Kauth made helpful comments. The study was carried out under the direction of R. R. Legault, J. D. Erickson, and R. F. Nalepka. The author gratefully acknowledges the help of all these co-workers.

## CONTENTS

1. Summary . . . . .	7
2. Introduction . . . . .	9
3. The Nine-Point Rules Defined . . . . .	11
3.1 The Nine-Point Likelihood Rule . . . . .	11
3.2 The Moving-Average Rule . . . . .	12
3.3 A Voting Rule . . . . .	13
3.4 Null Decisions and Boundary Detectors . . . . .	13
3.5 Other Multi-Element Rules . . . . .	14
4. Experimental Comparison of Nine-Point Rules . . . . .	16
5. Qualitative Comparison of Nine-Point Rules . . . . .	25
6. Conclusions and Recommendations . . . . .	35
6.1 Conclusions . . . . .	35
6.2 Recommendations . . . . .	35
Appendix A: How the Nine-Point Rules Are Programmed . . . . .	37
Appendix B: An Alternate Form of the Nine-Point Likelihood Criterion . . . . .	39
Appendix C: A Nine-Point Likelihood Model with Correlation . . . . .	40
References . . . . .	42
Distribution List . . . . .	43

PRECEDING PAGE BLANK NOT FILMED

## FIGURES

1. Percent Misclassified by Four Types of Decision Rules on Imperial Valley Fields Using the First Set of Training Fields . . . . .	22
2. Percent Misclassified by Four Types of Decision Rules on Imperial Valley Fields Using the Second Set of Training Fields . . . . .	23
3. Some Imperial Valley Fields Mapped by the One-Point Rule . . . . .	27
4. Some Imperial Valley Fields Mapped by the Nine-Point Likelihood Rule . . . . .	28
5. Some Imperial Valley Fields Mapped by the Best-7-of-9 Likelihood Rule . . . . .	29
6. Some Imperial Valley Fields Mapped by the Best-5-of-9 Likelihood Rule . . . . .	30
7. Some Imperial Valley Fields Mapped by the Moving Average Rule with the Biggest and Smallest Value in Each Channel Trimmed (TRIM = 1) . . . . .	31
8. Some Imperial Valley Fields Mapped by the Voting Rule . . . . .	32
9. The Best-7-OF-9 Likelihood Rule as a Boundary Detector . . . . .	33
10. The Voting Rule as a Boundary Detector . . . . .	34

## TABLES

1. Percent Misclassified by Four Types of Decision Rules on 20 Imperial Valley Training Fields Using the First Set of Training Fields . . . . .	18
2. Percent Misclassified by Four Types of Decision Rules on 22 Imperial Valley Test Fields Using the First Set of Training Fields . . . . .	19
3. Percent Misclassified by Four Types of Decision Rules on 20 Imperial Valley Training Fields Using the Second Set of Training Fields . . . . .	20
4. Percent Misclassified by Four Types of Decision Rules on 22 Imperial Valley Test Fields Using the Second Set of Training Fields . . . . .	21

# 1 SUMMARY

Nine-element rules decide what material to assign to a pixel on the basis of data from that pixel and from its eight immediate neighbors. They are applicable whenever a pixel is likely to represent the same material as its neighbors. The purpose of such rules is to gain recognition accuracy at only a slight extra cost in processing time. The consideration of neighboring data values adds some spatial information to what otherwise would be a purely multispectral decision process. Three such rules were implemented and tested:

The Nine-Point Likelihood Rule is the maximum likelihood decision rule derived from the assumption that the nine elements are an independent random sample from a multivariate normal distribution. It amounts to adding, for each material, the nine multivariate normal exponents and then choosing the material with the smallest sum. To prevent occasional alien points from disturbing the decision rule, we have modified it to sum only the  $m$  smallest exponents, where  $m = 1, \dots, 9$ .

The Voting Rule is applied after one-point decisions have been made on the nine pixels. It assigns to the center pixel the material most frequently recognized among the nine pixels. In case of a tie, the one-point decision on the center pixel is used.

The Moving Average Rule averages the nine data points and then applies the one-point rule. To lessen its sensitivity to alien points we have deleted the  $t$  largest and  $t$  smallest values of the nine in each channel, where  $t = 0, \dots, 4$ .

To compare and rank these three rules and the one-point rule, we ran a quantitative test by counting the number of points misclassified within each of 42 field interiors in the Imperial Valley, California. A result of the test was the following best-to-worst ranking of rule performance: nine-point likelihood rule with  $m = 9$ ; voting rule; moving average rule with  $t \neq 0$ ; moving average rule with  $t = 0$ ; one-point rule; and nine-point likelihood rule with  $m = 1$ . Performance of the nine-point likelihood rule improved steadily as  $m$  went from 1 to 9. For  $m = 9$ , its error rate was about one-half that of the one-point rule on the training sets, and on the test sets about three-fourths that of the one-point rule.

To supplement the results obtained on field interiors, we also made qualitative comparisons of maps generated by the different rules. To do this, we implemented an option to allow each rule to decide against all the alternative materials and display such decisions by leaving blanks on the map. Such null decisions create a white framework of roads, rivers, and other extraneous materials against which materials of interest stand out, thereby helping to produce a readable map.

The  $m=9$  rule with the null test splotted the map with white rectangles; this was because a single unusual point produces higher than normal exponent sums for a  $3 \times 3$  pixel rectangle around it. The rectangles disappeared for  $m = 7$ . Fine detail such as small roads seem to be lost by the nine-point rules. The null test for the voting rule (decide null if the winning vote total is too small) worked well in locating narrow boundaries distant from the material signatures but consistent with each other.

For some fields, the nine-point rules brought out an underlying pattern not readily apparent in a mixture of individual recognitions. For others, the nine-point rules seemed to find order where there was none. In either case, the contradictory character of the data was suppressed.

The null test can be used as a boundary detector by displaying each null point as a dark symbol and leaving everything else blank. Neither the  $m=7$  rule nor the voting rule succeeded well as a boundary detector. The  $m=7$  rule lost many small boundaries, and the voting rule lost the big extraneous areas.

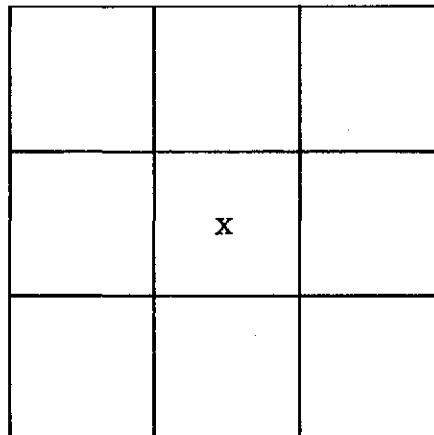
Our experiment comparing the nine-point rules and the one-point rule is based on but one data set; thus the conclusions from it are tentative, and the ultimate impact and utility of the nine-point approach have yet to be established. Because the nine-point rules performed successfully in the experiment, this suggests that they should be quantitatively and qualitatively tested on other data sets and encourages the implementation and comparison of other promising nine-point rules. There remains a need for development of a better boundary detector, one combining the principle of distance from known signatures with the principle of divided allegiance.



## 2 INTRODUCTION

The rules currently in use for multispectral recognition are single-element oriented—that is, they make a decision on each individual pixel without being influenced by decisions made on neighboring pixels. But for the many applications in which a pixel is likely to represent the same material as its immediate neighbors, a rule that takes neighboring data into account would be expected to perform better than a single-element rule.

Nine-element rules are designed to gain this advantage while preserving simplicity and speed. Such rules are applied in turn to each pixel of the scene in the context of its eight immediate neighbors arranged in a  $3 \times 3$  grid:



The rules assume that most or all of these nine pixels represent the same material, and they assign to the center pixel this majority material. Modest storage requirements and the small number of pixels playing a part in each decision make these rules practical.

Nine-element rules are most effective when the assumption of similarity of neighbors is most realistic. For this reason, one would expect nine-element rules to be more reliable than a single-element rule on the interiors of homogeneous areas and less precise on the boundaries. Nine-element rules would be applicable to data on agricultural fields collected at aircraft altitudes or in surveys of lakes and rivers; they would not be applicable, however, when the materials are "salted and peppered" across the scene, as in some geological data. When it is likely that neighboring pixels represent different materials, then it is also likely that many pixels represent more than one material. In this case, a mixture rule would be appropriate [1].

Although 25- and 49-element rules should not be ignored, we find them less attractive than nine-element rules because (1) they require storing five or seven scan lines at a time, thus taxing the fast-access storage of many computers; (2) each tier of pixels added to the group

makes the rule just that much more unclear as to field boundaries; (3) the increased number of pixels used slows down the rule; and (4) though an increase in accuracy can be expected in proceeding from one pixel to nine, a leveling off of accuracy occurs in going from 9 to 25 and from 25 to 49.

Another way of using data from neighboring pixels would be to define boundaries by a boundary-detection rule and then for each area enclosed to make a single decision applying to all pixels in the area [2]. (A generalization of one of the three decision rules defined in Sections 3.1 through 3.3 could be used.) This approach has certain difficulties: (1) human touch-up would be needed to fill gaps in the boundaries; (2) some data sets would not conform to the pattern of homogeneous areas surrounded by boundaries (as, for example, when water depth is mapped by multispectral recognition); and (3) if the shapes of homogeneous areas are more complicated than quadrilaterals, both a disk-storage system and a time-consuming algorithm would be needed in order to collect the data from a single field.

## THE NINE-POINT RULES DEFINED

### 3.1 NINE-POINT LIKELIHOOD RULE

The nine-point likelihood rule is a maximum likelihood decision rule based on an independent random sample of size nine from a normal distribution, rather than on a sample of size one from such a distribution. Increasing the sample size from one to nine will usually increase to a marked degree the accuracy of a statistical estimation procedure. For example, the accuracy of the sample mean as an estimate of the population mean is measured by the standard deviation of the sample mean. This quantity is a constant divided by the square root of the number of observations. Thus, the mean of a sample of nine observations would have one-third the standard deviation of a single observation.

Nine-point likelihood is simple to compute. It can be defined in terms of the one-point normal likelihood,

$$\text{constant } e^{-1/2 \left[ (x - \mu)^T R^{-1} (x - \mu) + \log_e |R| \right]}$$

where  $x$  is the data point

$\mu$  is the mean

$R$  is the covariance matrix of the distribution of the material under consideration

When the one-point rule is applied, only the quantity in the square brackets (hereinafter called the "exponent") is computed. The material producing the smallest exponent is the maximum likelihood choice. The nine-point likelihood, under the assumption of independence, is the product of the one-point likelihoods of the nine pixels. Hence, the nine-point maximum likelihood decision criterion is the sum of the nine exponents. The material with the smallest sum is the material chosen.

Computing each exponent is the most time-consuming task of the one-point decision rule. The nine-point likelihood rule, by comparison, does only the additional work of storing, retrieving, and summing the nine exponents. It can be efficiently applied by storing two unpacked scanlines of exponents and one packed scanline of data (see Appendix A). In short, from the standpoint of speed of execution and required storage, this rule is practical to apply.

Although it is unrealistic to assume that the nine points are independent, the rule derived from such an assumption may still be good. An analogous example is the one-point rule based on the normal distribution which worked well even on non-normal data [3]. The simplicity and practicality of the nine-point likelihood rule make it worth experimental trial, even if the full benefit one would expect to be derived from a valid model is not realized.

To guard against the possibility that not all the nine pixels represent the same distribution, the best-m-of-nine likelihood rule is also being studied. In this rule, we compute all nine exponents but sum only the m smallest. If one of the nine pixels includes roadways, a pile of rocks, or a patch of weeds, or if the data point has been garbled by the sensor, recorder, or digitizer, then the best-m-of-nine modification prevents such an abnormal point from smearing its own and neighboring recognitions.

This rule takes somewhat longer than the unmodified rule, because it requires sorting the nine exponents. Special cases, such as  $m = 8$ , could be programmed to run faster because one need only find the largest exponent and subtract it from the sum.

### 3.2 MOVING-AVERAGE RULE

The moving-average rule sums the nine data points and divides by nine to obtain an average data point for the  $3 \times 3$  grid; it then applies a single-element recognition rule. This rule is a common technique for reducing noise in the data. It and the nine-point likelihood rule are equally easy to apply. To give the moving-average rule the flexibility (similar to that of the best-m-of-nine likelihood rule) to reject odd points, we consider a trimmed mean rule. In every channel, the nine data values are ordered, the t largest and t smallest values are deleted, and the remaining values are averaged. When  $t = 0$ , the rule is an untrimmed moving-average rule; when  $t = 4$ , the median of the nine values in each channel is taken as the average data point.

Appendix B shows that the nine-point likelihood criterion can be expressed in the form

$$\log_e |R| + (\bar{X} - \mu)^T R^{-1} (\bar{X} - \mu) + \frac{1}{9} \sum_{i=1}^9 (X_i - \bar{X})^T R^{-1} (X_i - \bar{X})$$

where  $X_i$  is the i-th of the 9 data points

$\bar{X}$  is the mean of the nine points

$\mu$  is the mean of the material in question

R is its covariance matrix

If every material had the same covariance matrix, the last term would be the same for all materials and could be omitted. The first two terms comprise the moving-average criterion. Thus, if all materials have the same covariance matrix, the moving-average rule and the nine-point likelihood rule are identical in effect.

When the covariance matrices are unequal, however, the third term provides information about how closely the distribution within the nine pixels corresponds to the material covariance matrix, thereby helping in the recognition process.

Appendix C demonstrates that when the assumption of independence of the nine pixels is replaced by a simple correlation model, the maximum likelihood decision rule turns out to be a linear combination of the moving-average criterion and the nine-point likelihood criterion. The higher the correlation, the more weight is given the moving-average criterion and the less to the nine-point likelihood criterion.

### 3.3 VOTING RULE

The voting rule we studied is applied after one-point recognitions have been made on the nine pixels. The center pixel is assigned the material recognized most frequently among the nine. In case of tie, the one-point recognition on the center pixel is chosen. Its ease of application is about the same as that of the previously defined rules. Rules similar to the voting rule have been used to enhance space photographs and have been suggested for multispectral recognition.

### 3.4 NULL DECISIONS AND BOUNDARY DETECTORS

Recognition maps are made more readable if the category "none of these" is made part of the decision rule and printed as a blank. In one-element rules, the null decision is made when a point lies outside an equal-density ellipsoid of the winning signature, the size of which is so chosen that a point from the distribution has a prescribed probability (such as 0.001) of falling outside it. This test amounts to checking whether the quadratic form

$$(X - \mu)^T R^{-1} (X - \mu)$$

is greater than a constant  $C$  corresponding to the prescribed level.  $(X - \mu)^T R^{-1} (X - \mu)$  is the multivariate normal exponent without the  $\log_e |R|$  term. It has the chi-square distribution.  $C$  is the entry in the table of the chi-square distribution whose row number is the number of channels used and whose column heading is the significance level.

Although a predetermined level such as 0.001 is good for a start, the most readable map is usually obtained by trying several values of  $C$  and empirically obtaining the best value. To facilitate this search, we have separated the null test from the decision rule by writing a two-channel output tape; the first channel is the number of the winning signature and the second the value of the quadratic form.  $C$  becomes an input to the mapping program and several values may be tried efficiently.

For the best- $m$ -of-nine likelihood rule, the null criterion is the sum of the  $m$  smallest exponents minus  $m \log_e |R|$ . Under the assumption of independence, this criterion has the chi-square distribution with degrees of freedom equal to  $m$  times the number of channels. This criterion is written in the second channel of the output tape and a null decision is made

at the time of mapping. As with the one-point rule, a point failing the null test is on the outside of an equal-density ellipsoid chosen to reject legitimate points with a prescribed low level of probability.

For the voting rule, a null decision can be made whenever the winning vote total falls below a prescribed integer—a lack of consensus among the votes indicates a loss of confidence in the identification. The moving-average rule, which is a one-point rule applied to an average of the nine pixels, has the same null test as the one-point rule.

If the points failing the null test are mapped with a dark symbol and everything else left blank, the null test then becomes a boundary detector. The interiors of homogeneous areas corresponding to one of the given signatures would be left blank, outlined by pixels whose neighbors represent either more than one signature or some alien material whose signature was not provided.

The voting rule criterion would seem an appropriate boundary detector because a low winning vote total would indicate a divided allegiance in the neighborhood. We would expect the best- $m$ -of-nine criterion to be a better boundary detector for high values of  $m$  than for low values. If  $m$  were 7, for example, then three or more atypical pixels would significantly increase the boundary criterion; but if  $m$  were  $\leq 5$ , then a majority of the pixels would have to be atypical to produce such an increase. If a narrow boundary between homogeneous areas went through the middle of the  $3 \times 3$  grid, we would expect three or four pixels, but not a majority, to be atypical. Thus, the boundary would be detected by the  $m=7$  criterion but not by a  $m \leq 5$  criterion.

### 3.5 OTHER MULTI-ELEMENT RULES

Other promising multi-element processing rules can be defined, although we have not implemented and tested them. The three previously defined rules need not be restricted to a  $3 \times 3$  grid; they can be applied equally well to a  $5 \times 5$  or a  $7 \times 7$  grid or to an entire field. More complicated voting rules can be defined in which second choices are considered. The moving-average rule can be run with weights, the center element getting the most weight and the diagonal elements the least. A linear combination of the nine-point likelihood and moving-average decision criteria (shown in Appendix C to be equivalent to a nine-point likelihood rule based on a simple correlation model) could be implemented.

It has been suggested\* that the nine-point decision problem be treated as though it were a one-point decision problem with nine times as many channels, and that the fast but powerful linear decision rule [4] be employed. One would expect the nine-times-as-many-channels

---

\*In personal communication with R. J. Kauth and R. B. Crane of ERIM

rule to be more powerful and more time consuming than the three rules already presented. It might also be sensitive to the direction of flight over the training area. In agricultural applications, for example, it might be sensitive to the direction of the rows.

Because we have observed that between-field variation of a crop is different from variation observed within each field, a rule based on a between-field covariance matrix  $B$  and a within-field covariance matrix  $R$  would merit further study. One way to do this would be to use the nine-point likelihood rule in the form derived in Appendix B—that is, to choose the material  $j$  for which the expression

$$\text{constant}_j + (\bar{X} - \mu_j)^T R_j^{-1} (\bar{X} - \mu_j) + \frac{1}{9} \sum_{i=1}^9 (X_i - \bar{X})^T R_j^{-1} (X_i - \bar{X})$$

is smallest, except that  $B_j^{-1}$  replaces  $R_j^{-1}$  in the second term.

## EXPERIMENTAL COMPARISON OF NINE-POINT RULES

We have implemented the best-m-of-nine likelihood rule, the trimmed moving-average rule, and the voting rule by digital processing modules (described in Appendix A). To compare the effectiveness of each of these rules with the others and with the conventional one-point (quadratic) rule, we tried them on multispectral data collected from California's Imperial Valley (at 5000 ft in 1969). We chose these data for the experiment because we had confidence in the ground truth [5] and because some of the signatures were similar enough to make accurate recognition difficult, thereby offering us an opportunity to demonstrate differences in rule performance. The experiment was restricted to the 42 fields for which the ground truth was unequivocal and for which the scan angle was minimal. A previous study of the relative effectiveness of the quadratic and linear decision rules [4] has shown the error rate on these fields to be a sensitive measure of the power of the decision rule used.

Performance for each field was measured by the field error rate—that is, the number of elements misclassified divided by the number of elements in the field. So that the error rates would be comparable, we did not incorporate the null decision option in the rules. The crop error rates and the overall error rates were obtained by averaging the field error rates. The total rates were not computed by dividing the total number of misclassifications by the total number of points because that would have given too much weight to the results from the large fields. The overall error rate is estimated with two sources of error: the between-field variation and the within-field variation. Because we have found that the between-field variation overshadows the other and because the effect of between-field variation is minimized by an estimate giving each field equal weight, we have chosen that estimate.

The limits of the fields studied were defined as being several rows in from the apparent boundaries; this precaution excluded pixels on or near the boundaries which may have represented materials at variance with the ground truth. Thus, the experiment measured the performance of the rules on the interiors of fields and not at the boundaries. Because one expects the advantage of nine-point rules in the interiors to be offset somewhat by poorer performance on the boundaries, this was an unfortunate limitation but necessary since one cannot be sure of the ground truth of boundary pixels. Thus, the experimental results give an incomplete picture of rule performance unless they are interpreted side by side with qualitative results from the unabridged (i.e., field interiors plus boundaries) scene.

For each rule two computer runs were made, each with 20 training and 22 test fields, but with the training sets of the second run chosen from the test fields of the first. Later in this section, we report the results separately for training and test fields.



The rules tested were the one-point (quadratic) rule, the best- $m$ -of-nine likelihood rule ( $1 \leq m \leq 9$ ), the trimmed mean rule (with the number of values trimmed off each end varying from 0 to 4), and the voting rule. The nine-point rule with  $m = 1$  is not equivalent to the one-point rule because under the  $m=1$  rule, when the center element is closest to material A and one of the eight neighbors is closer still to material B, material B is chosen. The  $m=9$  rule is the original nine-point likelihood rule, the trim=0 rule is the usual moving-average rule, and the trim=4 rule is a moving-median rule. The one-point rule was applied to all pixels of the field except those on the edge; this permitted better comparison with the nine-point rules that were unable to classify edge pixels.

The results of the experiment are given in Tables 1 through 4 and Figs. 1-2. The figures illustrate the "totals" column of the tables. This column is the most important one in gauging the relative performance of the rules because the success of a rule with one crop may be more than offset by failures with other crops. And, according to the Bayesian theory of decisions, what counts is the minimization of total errors.

The four tables give the training field and test field results for the first and second choice of training fields. Looking first at the "totals" column of these tables, we see in all four cases a steady reduction in the percent misclassified by the best- $m$ -of-nine likelihood rule as  $m$  goes from 1 to 9. The one-point rule is better than the  $m=1$  rule in three cases and just as good in the fourth. The  $m=9$  rule, however, has one-half the error rate of the one-point rule on the training sets and three-fourths the one-point rate on the test sets. In all four cases, the  $m=9$  rule had lower rates than the voting rule or the trimmed mean rules. The voting rule and the trimmed mean rules performed substantially better than the one-point rule. The voting rule performed better than any trimmed mean rule in three cases and was about the same as the trimmed mean rules in the fourth. The only trend in the trimmed mean results is that the rule is uniformly a little worse when trim = 0 (untrimmed).

When we examine the columns of Tables 1-4 which show error rates for individual crops, however, the results are contradictory. With one exception, alfalfa, barley, and rye had decreasing error rates as  $m$  went from 1 to 9. In various columns of these four tables the decrease in percent misclassified is startling: 62 to 28 in one alfalfa column, 62 to 22 in a barley column, and 58 to 16 in a rye column. The three sugar beet columns and one lettuce column in which the numbers were large enough to discern a trend all had slightly increasing rates.

Is there any tendency for an upturn of rates at  $m = 9$ ? Looking at the 24 non-safflower results, we find  $m = 9$  worse than  $m = 8$  in 13 cases, better in 9 cases, and the same in 2. This is not a very significant trend and, in fact, disappears in the totalling.

TABLE 1. PERCENT MISCLASSIFIED BY FOUR TYPES OF DECISION RULES  
ON 20 IMPERIAL VALLEY TRAINING FIELDS USING THE FIRST SET OF  
TRAINING FIELDS

	<u>Total</u>	<u>Alfalfa</u>	<u>Barley</u>	<u>Lettuce</u>	<u>Rye</u>	<u>Bare Soil</u>	<u>Sugar Beets</u>	<u>Safflower</u>
One-Point Rule:	20.3	51.3	14.8	6.4	32.9	0.7	20.1	0
Best-m-of-Nine Likelihood Rule:								
m=1	20.0	51.4	13.7	1.0	57.7	0	7.2	0
m=2	17.6	49.0	8.7	0.9	50.5	0	5.8	0
m=3	16.1	47.5	6.2	0.4	45.2	0	5.5	0
m=4	14.9	46.3	4.6	0.3	39.2	0	5.3	0
m=5	14.1	45.5	3.7	0.1	35.1	0	5.1	0
m=6	13.3	43.7	3.2	0.2	30.7	0	5.4	0
m=7	12.7	42.3	3.5	0.2	25.9	0.3	5.6	0
m=8	12.0	40.2	3.8	0.2	21.5	0.4	6.3	0
m=9	11.7	39.7	3.5	1.3	16.3	1.0	7.5	0
Trimmed Mean Rule:								
trim=0	15.5	47.7	5.7	0.2	26.7	1.1	12.9	0
trim=1	14.1	46.1	4.4	0.1	25.2	0.3	9.8	0
trim=2	14.5	47.7	4.1	0.2	26.8	0	9.6	0
trim=3	13.7	46.4	3.8	0.2	26.3	0	7.0	0
trim=4	14.1	46.1	4.3	0.4	28.4	0	7.5	0
Voting Rule:	13.5	46.5	3.6	0.6	22.0	0	8.1	0

TABLE 2. PERCENT MISCLASSIFIED BY FOUR TYPES OF DECISION RULES  
ON 22 IMPERIAL VALLEY TEST FIELDS USING THE FIRST SET OF TRAIN-  
ING FIELDS

	<u>Total</u>	<u>Alfalfa</u>	<u>Barley</u>	<u>Lettuce</u>	<u>Rye</u>	<u>Bare Soil</u>	<u>Sugar Beets</u>	<u>Safflower</u>
One-Point Rule:	31.6	64.3	21.4	0.8	49.9	0.4	54.6	0.6
Best-m-of-Nine Likelihood Rule:								
m=1	34.1	69.5	22.7	0	78.4	0	44.3	0
m=2	31.6	68.7	16.8	0	66.1	0	45.4	0
m=3	30.0	65.8	15.1	0	58.3	0	46.7	0
m=4	28.6	63.5	13.3	0	51.3	0	47.3	0
m=5	27.3	61.4	11.9	0	45.4	0	47.8	0
m=6	26.5	58.8	11.8	0	41.6	0	48.7	0
m=7	25.5	55.8	11.5	0	37.8	0	49.4	0
m=8	25.0	54.1	12.3	0	34.0	0.1	50.2	0
m=9	24.8	51.1	13.4	0.8	31.2	0.5	52.6	0
Trimmed Mean Rule:								
trim=0	27.5	59.7	16.2	2.5	39.5	0.2	47.0	0
trim=1	26.2	57.7	14.4	0	38.5	0	46.7	0
trim=2	26.1	58.2	13.1	0	38.9	0	46.4	0
trim=3	26.2	59.2	12.3	0	39.0	0	46.9	0
trim=4	26.3	58.5	12.5	0	40.3	0	47.3	0
Voting Rule:	26.8	61.8	9.2	0	39.6	0	51.6	0

TABLE 3. PERCENT MISCLASSIFIED BY FOUR TYPES OF DECISION RULES  
ON 20 IMPERIAL VALLEY TRAINING FIELDS USING THE SECOND SET OF  
TRAINING FIELDS

	<u>Total</u>	<u>Alfalfa</u>	<u>Barley</u>	<u>Lettuce</u>	<u>Rye</u>	<u>Bare Soil</u>	<u>Sugar Beets</u>	<u>Safflower</u>
One-Point Rule:	21.8	46.1	43.1	1.3	18.4	0.4	12.5	0.2
Best-m-of-Nine Likelihood Rule:								
m=1	26.1	61.8	62.1	0	11.7	0	1.3	0
m=2	24.2	60.3	55.6	0	9.3	0	0.6	0
m=3	21.9	55.3	49.6	0	8.5	0	0.5	0
m=4	19.5	50.3	42.7	0	8.3	0	0.5	0
m=5	17.6	46.6	36.9	0	8.2	0	0.5	0
m=6	15.6	41.8	32.1	0	7.7	0	0.6	0
m=7	13.3	36.7	25.5	0	7.5	0	0.9	0
m=8	12.3	33.1	23.8	0	7.6	0.2	1.0	0
m=9	11.4	28.2	21.9	0.6	7.5	0.8	2.6	0
Trimmed Mean Rule:								
trim=0	18.9	42.6	46.4	0	6.5	1.1	1.8	0
trim=1	17.6	39.5	46.0	0	5.3	0	0.1	0
trim=2	17.6	39.4	46.0	0	5.2	0	0.1	0
trim=3	17.6	39.5	45.5	0	5.4	0	0.2	0
trim=4	17.2	37.7	44.7	0	6.6	0	0.4	0
Voting Rule:	14.2	31.9	35.0	0	7.5	0	0.7	0

TABLE 4. PERCENT MISCLASSIFIED BY FOUR TYPES OF DECISION RULES  
ON 22 IMPERIAL VALLEY TEST FIELDS USING THE SECOND SET OF  
TRAINING FIELDS

	<u>Total</u>	<u>Alfalfa</u>	<u>Barley</u>	<u>Lettuce</u>	<u>Rye</u>	<u>Bare Soil</u>	<u>Sugar Beets</u>	<u>Safflower</u>
One-Point Rule:	32.8	51.8	33.5	21.0	64.1	1.1	40.2	0
Best-m-of-Nine Likelihood Rule:								
m=1	35.7	68.7	48.8	11.5	55.8	0	20.9	0
m=2	33.4	64.9	44.0	11.8	55.3	0	18.7	0
m=3	32.1	62.1	40.4	12.1	55.3	0	19.4	0
m=4	30.4	58.4	36.1	12.6	56.2	0	19.8	0
m=5	28.8	54.4	32.0	12.8	56.9	0	20.6	0
m=6	27.2	50.9	27.7	13.1	57.3	0	21.7	0
m=7	25.6	47.8	22.2	13.8	57.6	0.3	23.1	0
m=8	24.1	44.9	18.0	13.5	57.7	0.4	23.6	0
m=9	23.5	43.2	15.3	13.8	57.8	0.9	25.9	0
Trimmed Mean Rule:								
trim=0	29.9	50.4	36.9	12.7	54.7	3.4	24.2	0
trim=1	28.9	49.8	37.4	13.0	53.7	0.3	21.6	0
trim=2	28.6	49.5	36.9	12.6	53.7	0	21.3	0
trim=3	28.4	49.4	36.4	13.3	54.3	0	20.4	0
trim=4	28.3	48.9	35.3	14.1	55.6	0	21.2	0
Voting Rule:	26.7	46.2	27.6	14.4	55.9	0	26.4	0

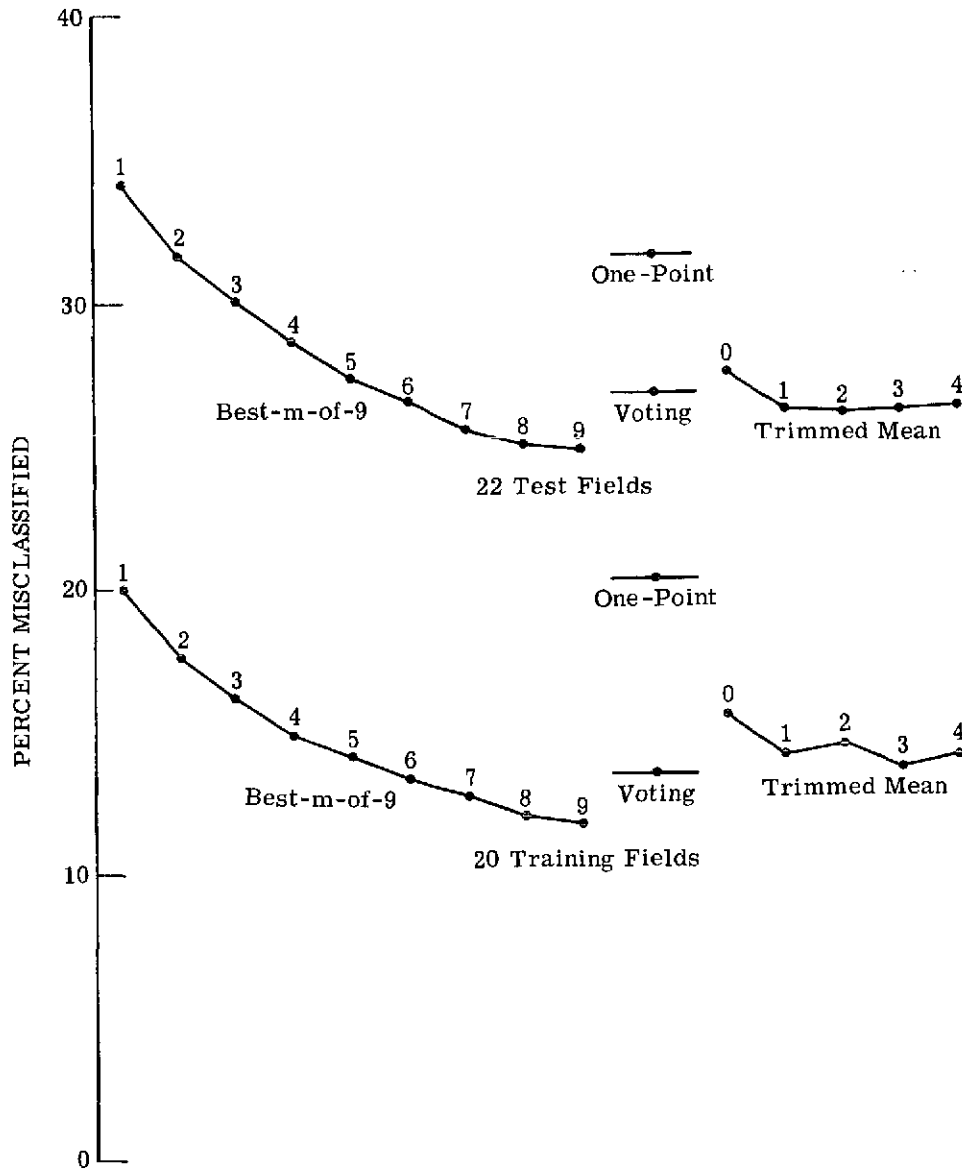


FIGURE 1. PERCENT MISCLASSIFIED BY FOUR TYPES OF DECISION RULES ON IMPERIAL VALLEY FIELDS USING THE FIRST SET OF TRAINING FIELDS

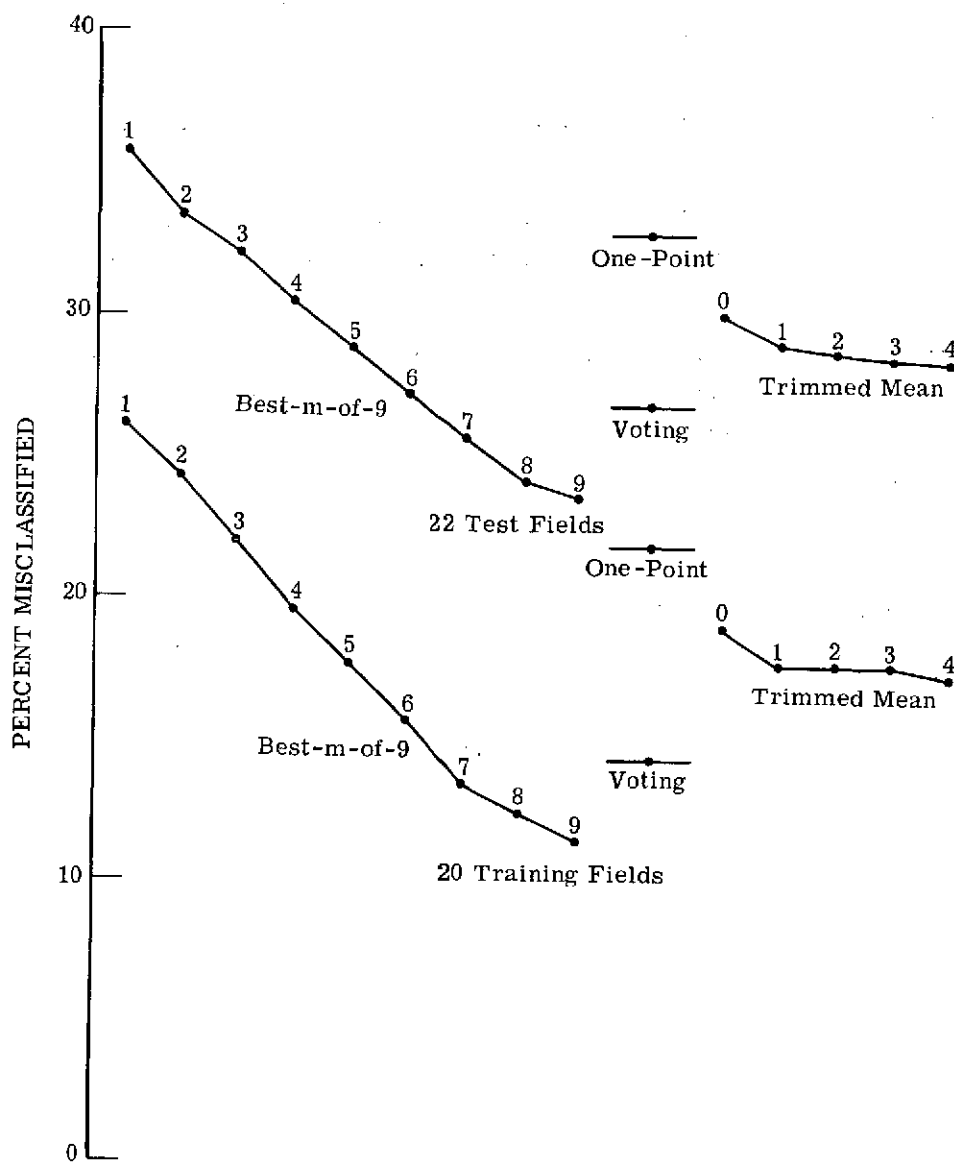


FIGURE 2. PERCENT MISCLASSIFIED BY FOUR TYPES OF DECISION RULES ON IMPERIAL VALLEY FIELDS USING THE SECOND SET OF TRAINING FIELDS

If we compare the one-point rule, the  $m=9$  rule, the  $\text{trim}=1$  rule, and the voting rule for the 17 cases in which the numbers are substantial, we discover the following: the one-point rule is the worst one in every case but one, the  $m=9$  rule is best in 10 cases, the  $\text{trim}=1$  rule best in 6 cases, and the voting rule best once. Comparing just the voting rule and the  $\text{trim}=1$  rule, it's an 8-to-9 split.

The slight inferiority of the untrimmed to the  $\text{trim}=1$  moving-average rule is consistent in crop error rates. Of 23 non-zero cases, the untrimmed rule did best only twice and equally well once.

Since total error rate is a reasonable measure of performance, we summarize the results of the experiment as follows: For the interiors of the homogeneous areas tested, the order of performance of the rules from best to worst is

- (1) nine-point likelihood
- (2) voting
- (3) trimmed mean
- (4) untrimmed mean
- (5) one-point
- (6) best-likelihood-of-nine

The error rate of the best rule is one-half that of the one-point rule on the training sets and three-fourths that of the one-point rule on the test sets.



## QUALITATIVE COMPARISON OF NINE-POINT RULES

To supplement these quantitative results, we used the rules previously described to make maps of a stretch of the Imperial Valley based on data containing many of the fields appearing in the quantitative study. These maps, included as Figs. 3 through 8, show the results of using the one-point rule, the nine-point likelihood rule with  $m = 9$ ,  $m = 7$ , and  $m = 5$ , the moving average rule with trim = 1, and the voting rule, respectively.

Implementing an option to allow each rule to decide against all alternative materials (see Section 3.4), we allowed such null decisions to be displayed in the form of blanks on the map. Such decisions leave a white framework of roads, rivers, and other extraneous materials, against which materials of interest show up, thereby helping produce a readable map.

The  $m=9$  rule, the best one in the quantitative study, has an unfortunate tendency to splotch the map with white rectangles (indicative of lower probability density), even when the null test limit is set higher than normal. This is because a single unusual point produces higher than normal exponent sums for a  $3 \times 3$  rectangle surrounding it. The rectangles disappear, however, when  $m = 7$ . The  $m=9$  rule widens big roads and, as  $m$  drops from 9 to 5, the rules show an increasing tendency to lose sight of small roads. The voting rule does a fairly good job of picking up small roads but tends to fill in wide ones.

For many fields, the one-point rule reports a "tossed salad" of recognitions, making it difficult to perceive the basic pattern. The nine-point rules make it easier to perceive order in the recognitions, but they have a tendency to find order where there is none. A field of sugar beets and one of alfalfa, disguised on the one-point map by a scattering of false recognitions of other crops, are accurately displayed on the nine-point maps. Another field in the one-point map, appearing to be sugar beets but producing many contradictory recognitions, is smoothed out by the nine-point maps to display nearly pure sugar beets. According to ground truth, however, it is a barley field. The doubtfulness of the one-point recognition comprised important information that was lost by the nine-point maps. In addition, a patch of weeds, which on the one-point map look like nothing but a shapeless mixture, is defined on the nine-point likelihood maps as a rye field.

Figures 9 and 10 show the use of the  $m=7$  rule and the voting rule, respectively, as boundary detectors. The results are not impressive. Of course, the data are not clear-cut; rather, they were chosen to present a challenge to the decision rules. Even so, each rule exhibits a deficiency. The  $m=7$  rule loses boundaries other than large roads because its recognition of a boundary point requires a larger than usual sum of distances from the chosen signature. The voting rule reports false recognitions on large roads and other areas not associated with one of the material signatures but consistent in signal with them. In such instances, whatever

distant signature happens to be preferred is likely to pull a majority of votes. An example of this tendency is a pastured field recognized by the voting rule as a field of lettuce.

A boundary detector that combines the principle of distance with that of divided allegiance would probably work better than either of the methods presented.

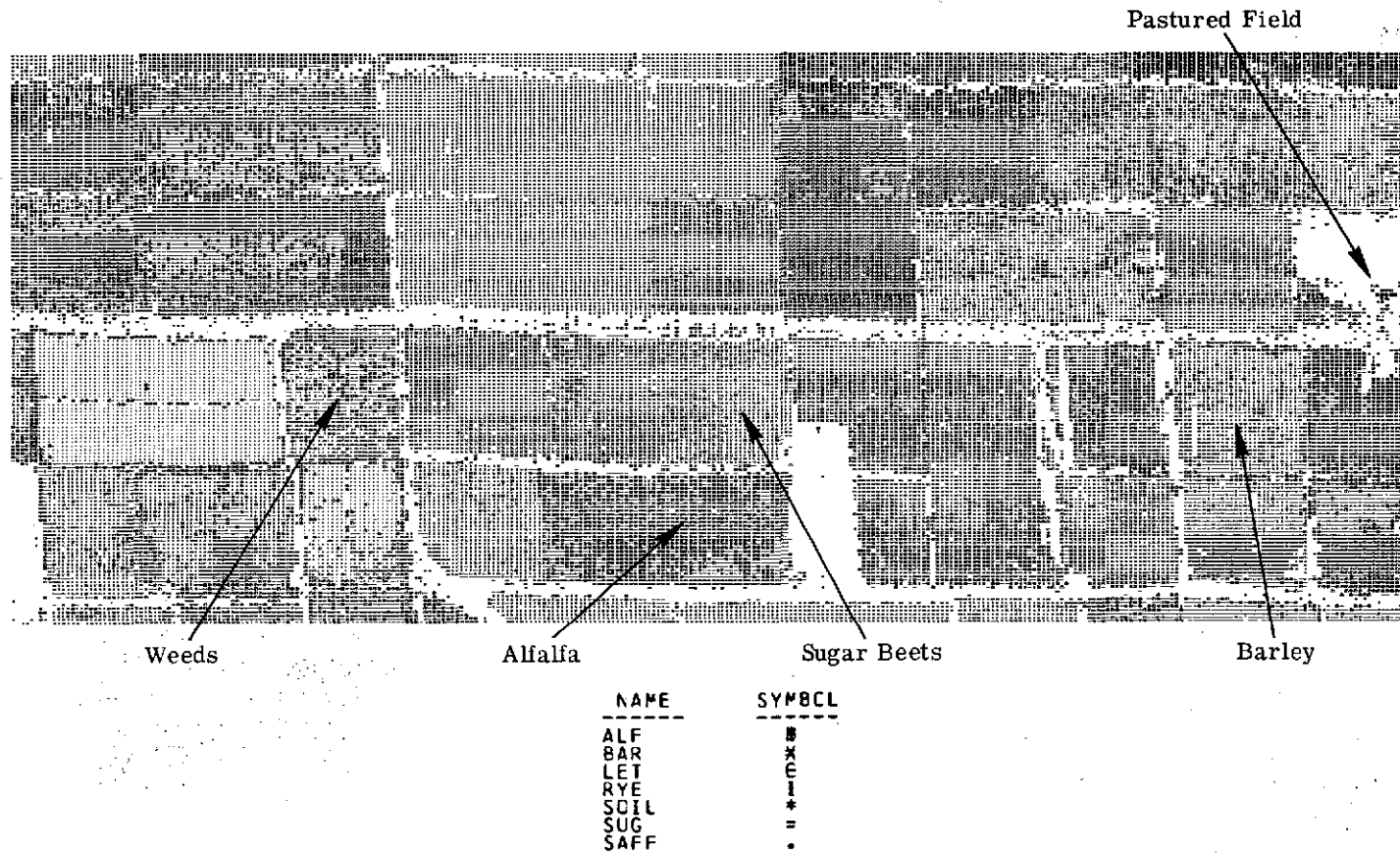


FIGURE 3. SOME IMPERIAL VALLEY FIELDS MAPPED BY THE ONE-POINT RULE

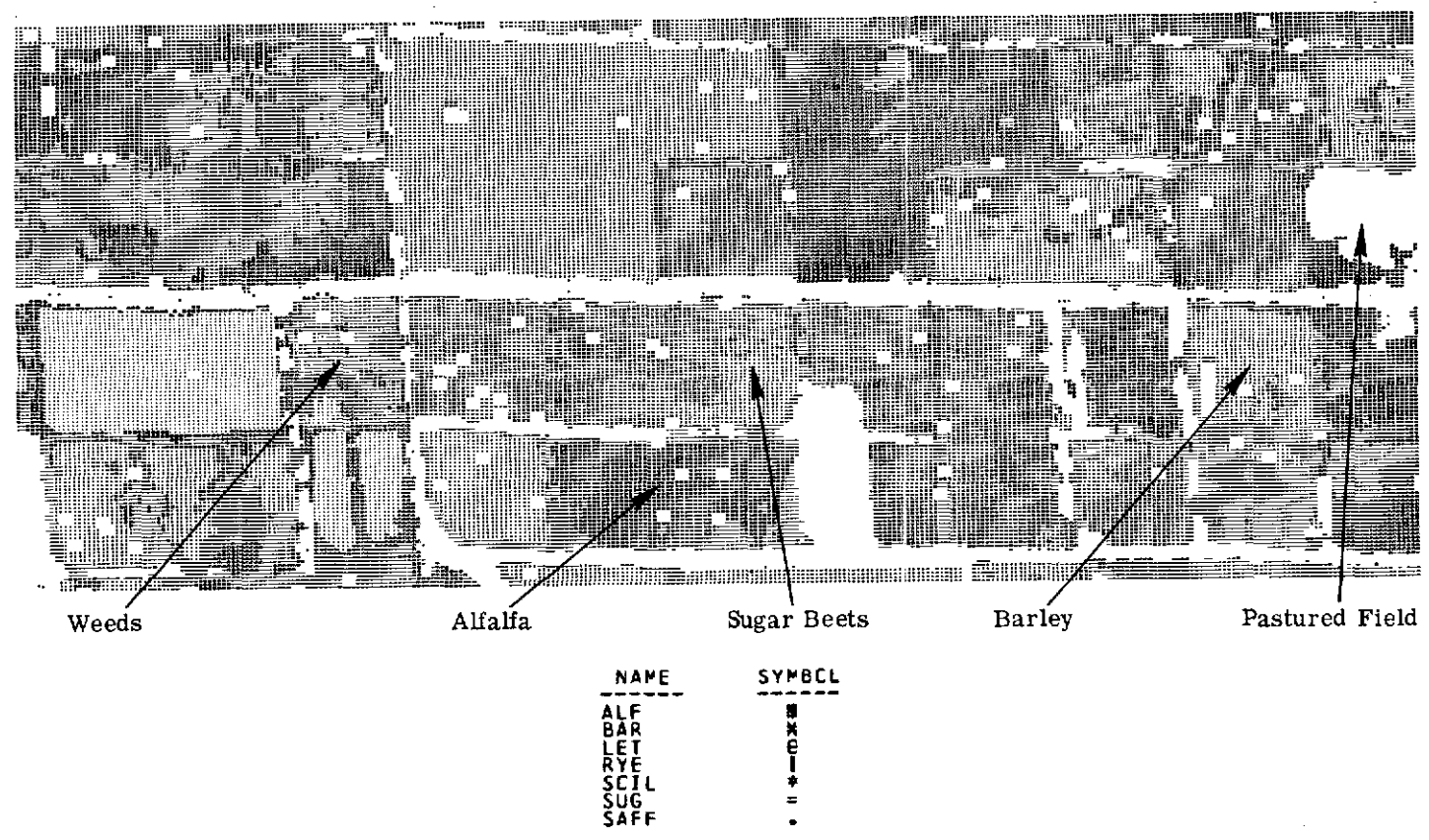


FIGURE 4. SOME IMPERIAL VALLEY FIELDS MAPPED BY THE NINE-POINT LIKELIHOOD RULE

29

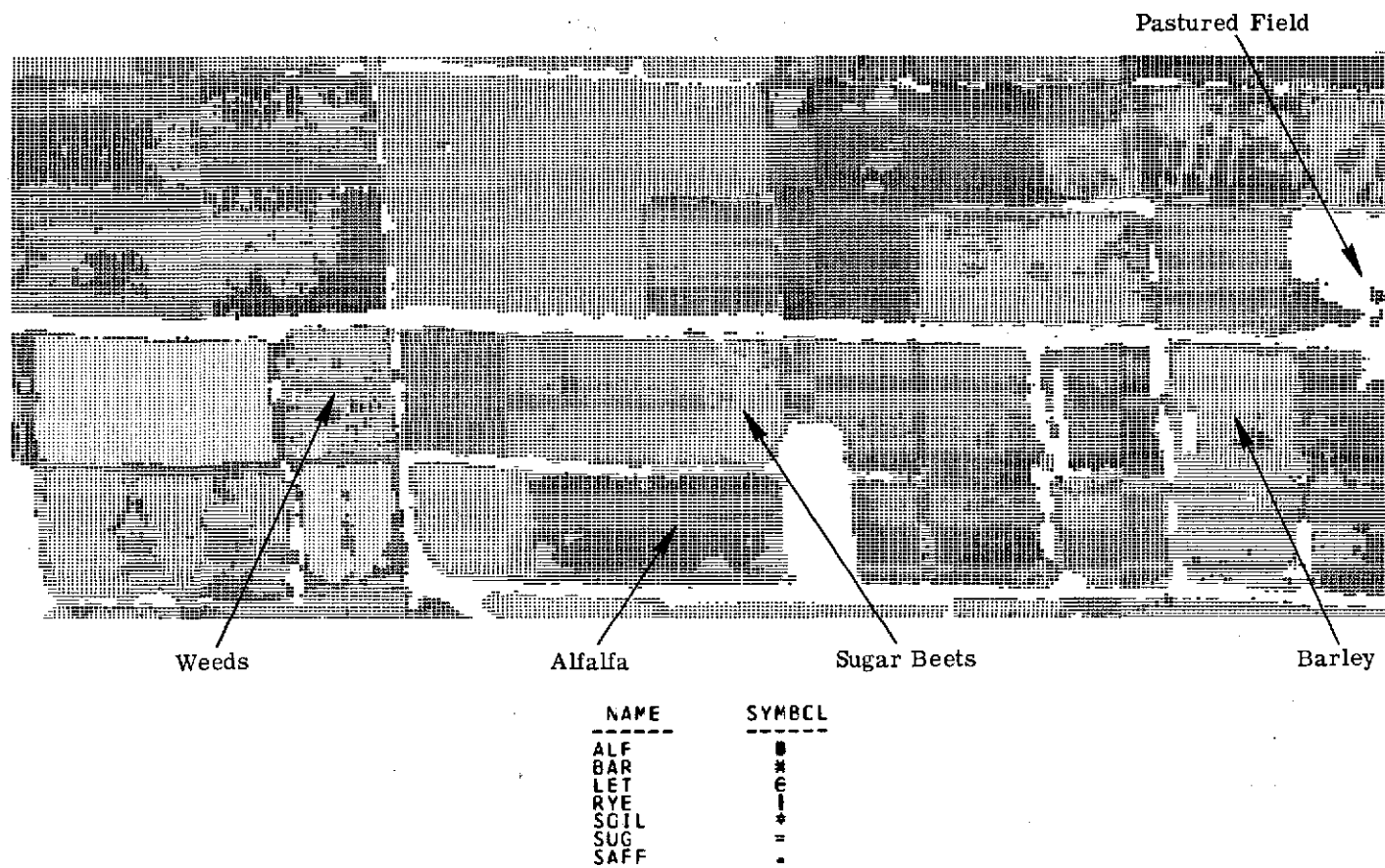


FIGURE 5. SOME IMPERIAL VALLEY FIELDS MAPPED BY THE BEST-7-OF-9 LIKELIHOOD RULE

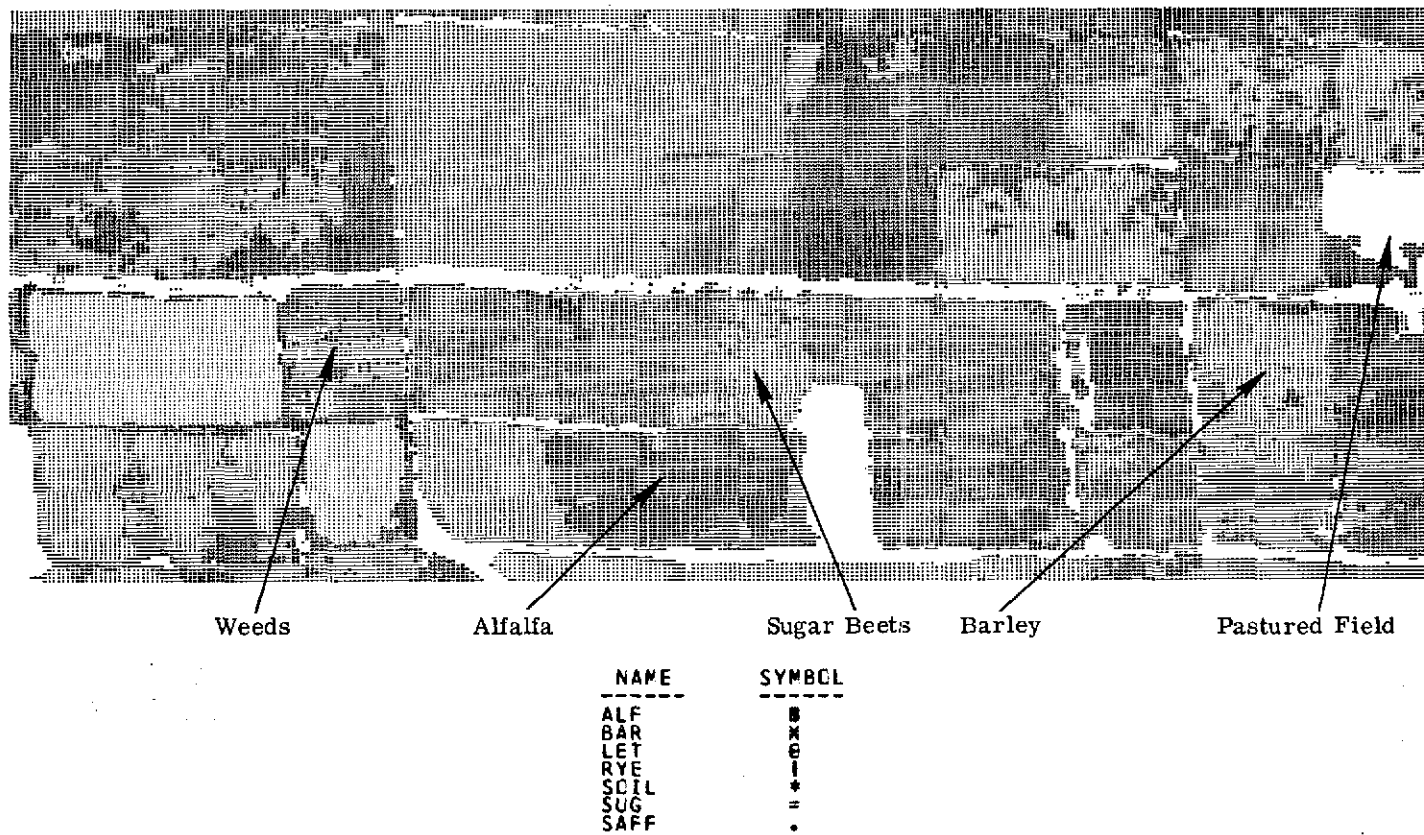


FIGURE 6. SOME IMPERIAL VALLEY FIELDS MAPPED BY THE BEST-5-OF-9 LIKELIHOOD RULE

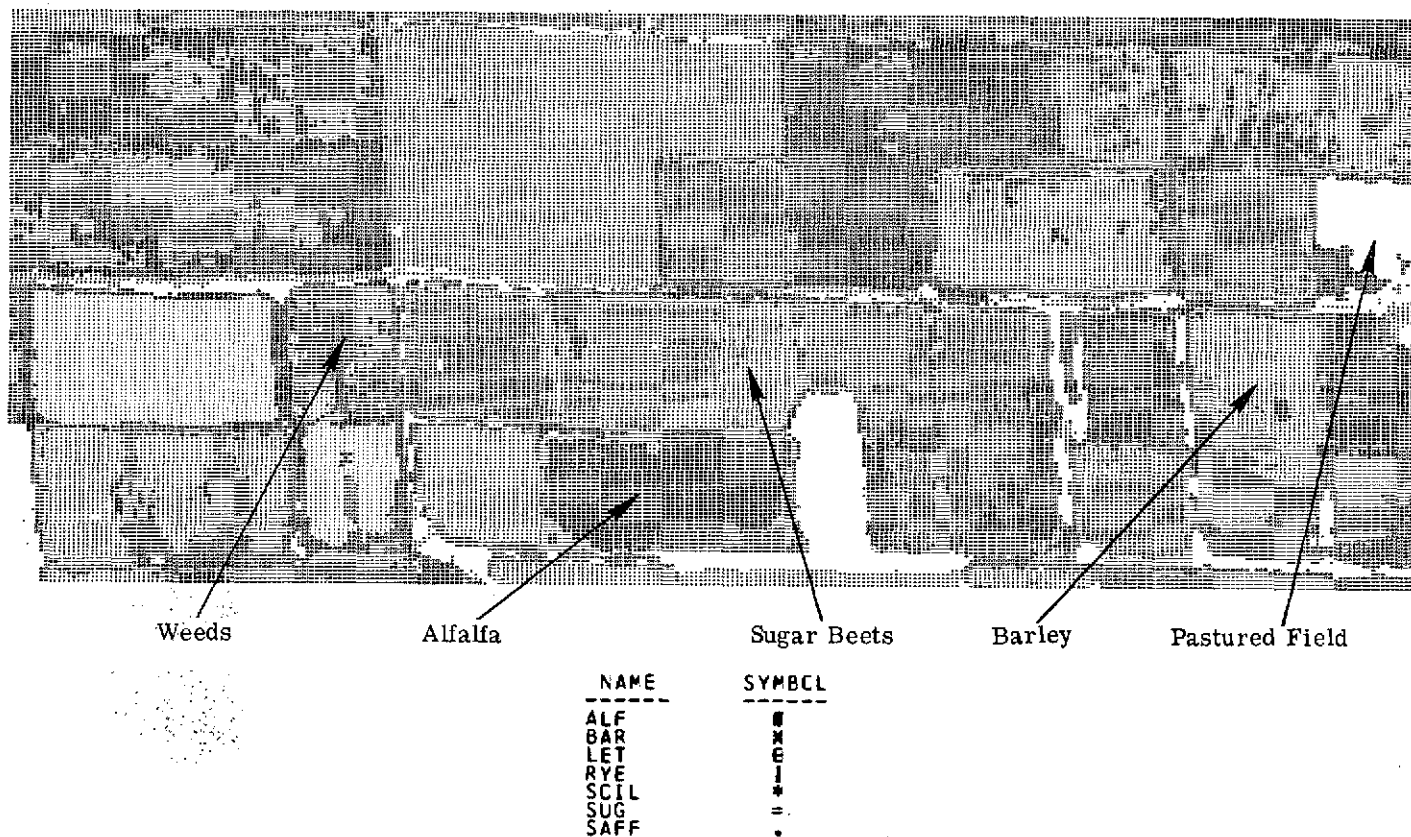


FIGURE 7. SOME IMPERIAL VALLEY FIELDS MAPPED BY THE MOVING AVERAGE RULE WITH THE BIGGEST AND SMALLEST VALUE IN EACH CHANNEL TRIMMED (TRIM = 1)

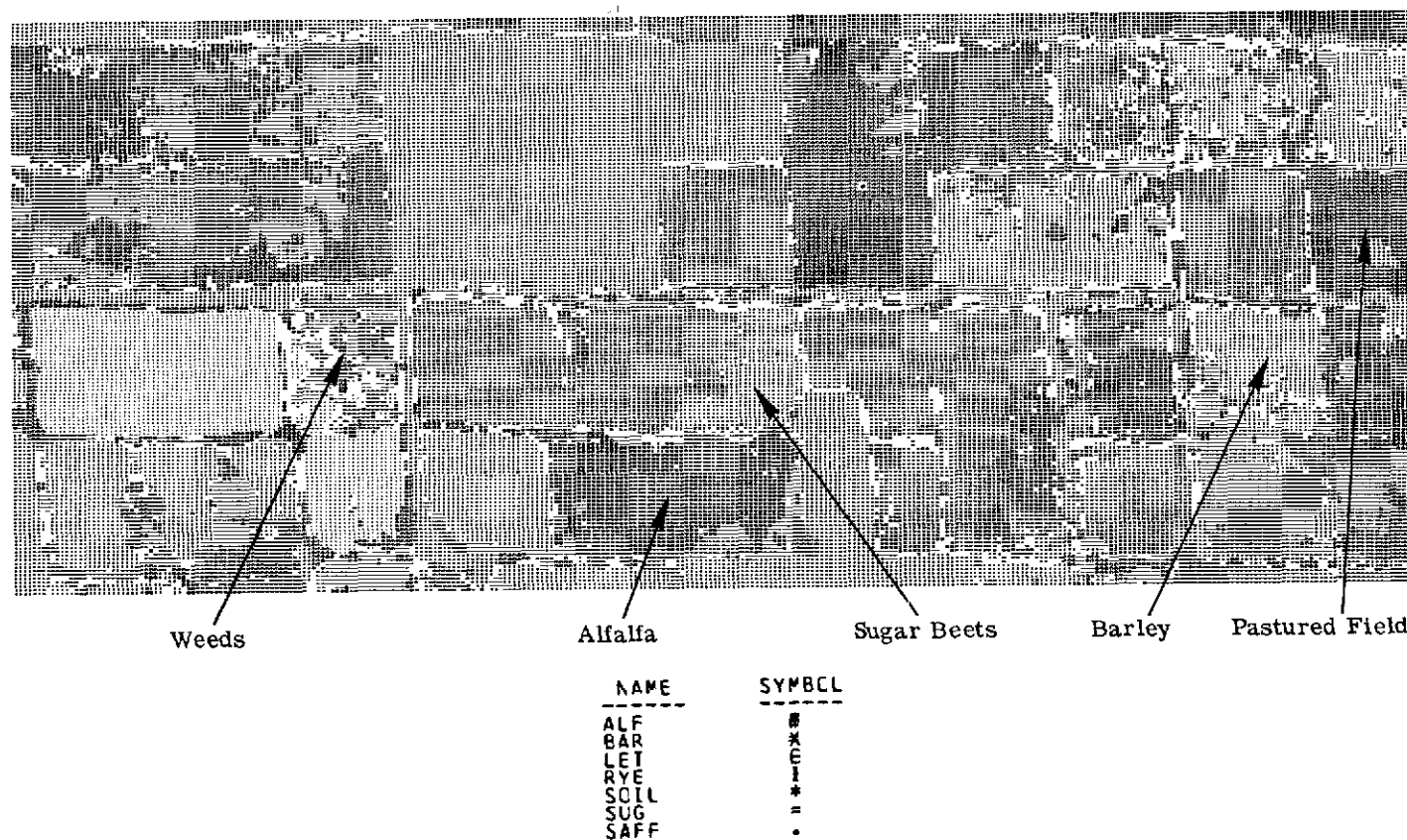


FIGURE 8. SOME IMPERIAL VALLEY FIELDS MAPPED BY THE VOTING RULE



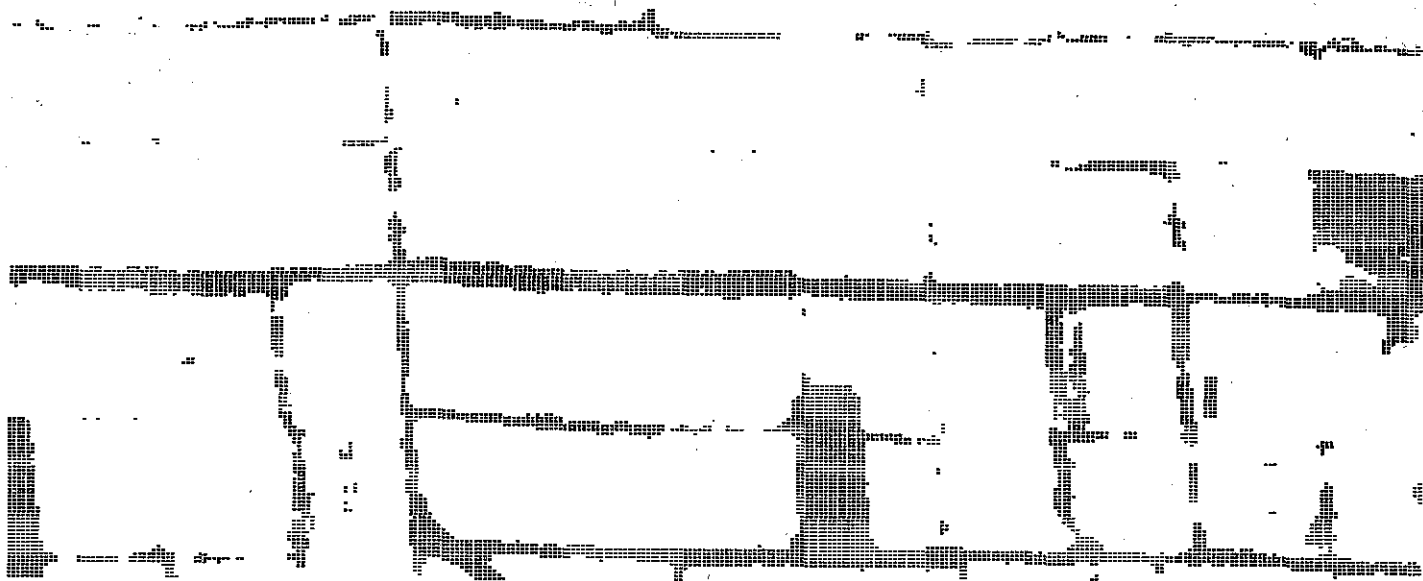


FIGURE 9. THE BEST-7-OF-9 LIKELIHOOD RULE AS A BOUNDARY DETECTOR



FIGURE 10. THE VOTING RULE AS A BOUNDARY DETECTOR

## 6

## CONCLUSIONS AND RECOMMENDATIONS

## 6.1 CONCLUSIONS

The experiment comparing the nine-point rules and the one-point rule is based on one data set. Therefore, the conclusions that follow are tentative. The ultimate impact and utility of the nine-point approach has yet to be established.

The nine-point likelihood rule, modified to sum the best seven of nine exponents, shows promise as a recognition rule to increase accuracy. While the unmodified nine-point likelihood rule performs best on field interiors, it is unsatisfactory when used with a null test to make a recognition map because of its tendency to expand deviant pixels into  $3 \times 3$  blank areas. On field interiors, the moving-average rule and voting rules perform better than the one-point rule. The moving-average rule does a little better, even on field interiors, when the largest and smallest values in each channel are deleted from the sum. Preliminary qualitative results indicate that the nine-point rules are less precise than the one-point rule in recognizing fine structure in a scene, which indicates that their most useful application is to scenes consisting mostly of large, homogeneous areas. The voting rule and best-seven-of-nine rule are not very satisfactory boundary detectors.

## 6.2 RECOMMENDATIONS

Because the nine-point rules studied performed successfully, they should be quantitatively and qualitatively tested on other data sets for which good ground truth exists. This performance also encourages the implementation and comparison of other nine-point rules, such as those mentioned in Section 3.5. One of the more promising of these rules treats the nine-point decision problem as a one-point decision problem with nine times as many channels. Another uses two covariance matrices, one for between-field variation (which is used with the mean of the nine pixels) and the other for within-field variation (used with the local variation among the nine pixels). The development of a better boundary detector, combining the principle of distance from known signatures with the principle of divided allegiance, is indicated.

## Appendix A HOW THE NINE-POINT RULES ARE PROGRAMMED

The Institute has a multispectral subsystem of the software system called ERIMS that provides the following: mounting, reading, and unpacking of data tapes; calling of modules to process the data; packing of output data values, four to a word; and writing of an output tape.

The subsystem consists of subroutines PROCESS and POINT. At the point-processing stage, a module accepts an input data point called DATUM, consisting of NCHAN channel values. The module modifies the DATUM vector in some way, storing the output vector in DATUM. After all the prescribed modules have been called, POINT and PROCESS pack up DATUM, adding it on to the output line that will be written on tape. If several operations are to be performed, they can be done as separate jobs (with the intermediate tape for one job providing the input for the next) or they all can be run together (with each module picking up the output DATUM vector from the previous module). The modules are also called—at an earlier stage when initial calculations are made, and at a later stage for final calculations and printing of results.

The m-fold rule is carried out by two modules. The first, DENS (short for DENSITY), finds in DATUM the channel values of a multispectral data point and calculates, for each signature read, the multivariate exponent of that data point; it then stores this result in DATUM. Thus, DATUM has NCHAN values coming in and NSIG values going out, where NSIG is the number of signatures.

The second module, LIKE9, picks up the nine relevant DATUM vectors by calling an assembly-language subroutine SAVE9 that is used by all the nine-point rules. SAVE9 stores two unpacked lines of DATUM vectors in the auxiliary core memory of our IBM 7094 computer, retrieves the DATUM vectors of the  $3 \times 3$  grid, and stores them in an  $NSIG \times 3 \times 3$  array DAT9. Only two lines need be stored because the third is the one being unpacked point by point. This most current line replaces, point by point, the least current one in auxiliary memory. For example, suppose you have just finished with point 30. One line in auxiliary memory consists of the current line through point 30 and the least current line from point 31 to the end. The other stored line is the second most current line.

LIKE9 works with the DAT9 array of nine DATUM vectors, each a vector of NSIG exponents. For each channel (i. e., for each signature), LIKE9 sorts the nine exponents and sums the m smallest—m is an input to LIKE9 in the initialization stage. The number of the channel with the smallest sum is put out as DATUM(1) and the value of the sum, appropriately scaled, as DATUM(2).

LIKE9 does nothing but store data points for the first two lines or for the first two points of each line, so that when it does become active, the DAT9 array contains nine meaningful data

points. After processing line number 3, it puts out a line number 2—because the calculation concerns the center point of the  $3 \times 3$  grid. The point numbering is analogous. Thus, a line is lost from top and bottom of each run, as well as a point from the beginning and end of each line.

Alone, LIKE9 produces a tape but no directly interpretable output. Used with a mapping module (either as a single job or with an intermediate tape), however, it produces a recognition map. The mapping module can be set to print a blank whenever the second channel (the sum of exponents) gets too large; the result is that all pixels distant from any input signature are left blank.

LIKE9 can also be used with the module TALLY to count up recognitions within the rectangle specified, print the count at the end of the run, and punch a card with the same information. The cards can then be read by program DISPLAY to print out misclassification rates for each field, for each crop, for all training sets, and for all test sets. One vector giving ground truth and another identifying the training sets in the deck of field cards are needed as inputs for DISPLAY.

The moving-average rule is carried out by two modules, AVE9 and QRULE. In the initialization stage, AVE9 reads an integer TRIM that must be between 0 and 4. AVE9 reads the original data tape, using subroutine SAVE9 to give it the nine relevant points in the array DAT9. For each channel I, AVE9 orders the nine values, deletes the TRIM largest and TRIM smallest, sums the rest, divides by the number summed, and then puts that number into DATUM(I). The effect of AVE9 is to replace each data point by an averaged data point.

QRULE is the one-point maximum likelihood decision rule. It reads each data point, computes the exponent for each signature, then puts the number of the signature with the smallest exponent in DATUM(1), and the value of that exponent in DATUM(2). QRULE can be followed either by TALLY or a mapping module. Though usually used with original data, it can just as easily accept the average data points put out by AVE9.

The voting rule is carried out by the modules QRULE and VOTE9. QRULE supplies the recognition (i.e., the winning signature number) in DATUM(1). VOTE9 uses subroutine SAVE9 to store the nine relevant recognitions in the array DAT9. It goes through the nine, tallying the number of recognitions of each signature. Then, the number of the signature with the most recognitions is put into DATUM(1) and the winning vote total in DATUM(2). In case of tie, the signature number of the center pixel is put into DATUM(1). VOTE9 can be followed either by TALLY or a mapping program.

## Appendix B AN ALTERNATE FORM OF THE NINE-POINT LIKELIHOOD CRITERION

The nine-point likelihood criterion is

$$\sum_{i=1}^9 \left[ (X_i - \mu)^T R^{-1} (X_i - \mu) + \log_e |R| \right]$$

where vector  $X_i$  is point number  $i$  of the nine points

$\mu$  is the mean vector of the material being considered

$R$  is the covariance matrix of this material

The material minimizing this criterion is the one chosen.

Dividing by 9 and adding and subtracting the mean  $\bar{X}$  of the nine points, the criterion becomes

$$\begin{aligned} \log_e |R| + \frac{1}{9} \sum_{i=1}^9 (X_i - \bar{X} + \bar{X} - \mu)^T R^{-1} (X_i - \bar{X} + \bar{X} - \mu) = \log_e |R| \\ + \frac{1}{9} \left[ \sum_{i=1}^9 (X_i - \bar{X})^T R^{-1} (X_i - \bar{X}) + \sum_{i=1}^9 (\bar{X} - \mu)^T R^{-1} (\bar{X} - \mu) \right. \\ \left. + 2 \sum_{i=1}^9 (X_i - \bar{X})^T R^{-1} (\bar{X} - \mu) \right] \end{aligned}$$

$\bar{X}$ ,  $\mu$ , and  $R$  stay constant for  $i = 1, \dots, 9$ . The last term,

$$2 \left[ \sum_{i=1}^9 (X_i - \bar{X})^T \right] R^{-1} (\bar{X} - \mu) = 0$$

because the sum of deviations from the mean is 0. So the criterion is

$$\log_e |R| + (\bar{X} - \mu)^T R^{-1} (\bar{X} - \mu) + \frac{1}{9} \sum_{i=1}^9 (X_i - \bar{X})^T R^{-1} (X_i - \bar{X})$$

which is the moving-average criterion plus a term measuring how closely the variation among the nine points is in accordance with the material covariance matrix.

### Appendix C A NINE-POINT LIKELIHOOD MODEL WITH CORRELATION

We will derive a nine-point likelihood criterion based on a simple correlation model. Let  $n$  be the number of channels. We consider the nine points  $X_1, \dots, X_9$  to be a single point  $X^*$  with nine times as many channels:

$$X^* = X_{11}, \dots, X_{1n}, X_{21}, \dots, X_{2n}, \dots, X_{91}, \dots, X_{9n}$$

We assume that its covariance matrix is of the form

$$R^* = \begin{bmatrix} R & \rho R & \rho R & \dots & \rho R \\ \rho R & R & \rho R & \dots & \rho R \\ \rho R & \rho R & R & \dots & \rho R \\ \dots & \dots & \dots & \dots & \dots \\ \rho R & \rho R & \rho R & \dots & R \end{bmatrix}$$

$R^*$  is a  $9 \times 9$  matrix of  $n \times n$  matrices. In other words, this simple model assumes that the correlation between any two points in the  $3 \times 3$  grid is  $\rho$ . To find the covariance of channel  $j$  of one point and channel  $k$  of another, multiply the single-point covariance  $R_{jk}$  by  $\rho$ .

The one-point maximum likelihood criterion applied to the super point  $X^*$  is

$$\log_e |R^*| + (X^* - \mu^*)^T R^{*-1} (X^* - \mu^*)$$

Written out in detail without the  $\log_e R^*$  term, it is

$$\begin{aligned} & \sum_{i=1}^9 (X_i - \mu)^T R^{-1} (X_i - \mu) + \rho \sum_{i \neq j} (X_i - \mu)^T R^{-1} (X_j - \mu) \\ &= \rho \sum_{i=1}^9 \sum_{j=1}^9 (X_i - \mu)^T R^{-1} (X_j - \mu) + (1 - \rho) \sum_{i=1}^9 (X_i - \mu)^T R^{-1} (X_i - \mu) \\ &= \rho \sum_{i=1}^9 (X_i - \mu)^T R^{-1} \sum_{j=1}^9 (X_j - \mu) + \dots \\ &= 9\rho \sum_{i=1}^9 (X_i - \mu)^T R^{-1} (\bar{X} - \mu) + \dots \\ &= 9\rho \left[ \sum_{i=1}^9 (X_i - \mu)^T \right] R^{-1} (\bar{X} - \mu) + \dots \\ &= 81\rho (\bar{X} - \mu)^T R^{-1} (\bar{X} - \mu) + \dots \end{aligned}$$

So, the criterion is

$$\log_e |R^*| + 81\rho(\bar{X} - \mu)^T R^{-1}(\bar{X} - \mu) + (1 - \rho) \sum_{i=1}^9 (X_i - \mu)^T R^{-1}(X_i - \mu)$$

which is a linear combination of the moving-average criterion and the nine-point likelihood criterion.

This can be put into another form by using the results derived in Appendix B. The last term becomes

$$9(1 - \rho)(\bar{X} - \mu)^T R^{-1}(\bar{X} - \mu) + (1 - \rho) \sum_{i=1}^9 (X_i - \bar{X})^T R^{-1}(X_i - \bar{X})$$

Thus, the criterion is

$$\log_e |R^*| + (9 + 72\rho)(\bar{X} - \mu)^T R^{-1}(\bar{X} - \mu) + (1 - \rho) \sum_{i=1}^9 (X_i - \bar{X})^T R^{-1}(X_i - \bar{X})$$



## REFERENCES

1. H. M. Horwitz, P. D. Hyde, and W. Richardson, Improvements in Estimating Proportions of Objects from Multispectral Data, Technical Report 190100-25-T, Environmental Research Institute of Michigan, Ann Arbor. (In Publication.)
2. A. G. Wacker and D. A. Landgrebe, The Minimum Distance Approach to Classification. Note 100771, Laboratory for Applications of Remote Sensing Information, West Lafayette, Ind., 1971, pp. 127-128.
3. R. B. Crane, W. A. Malila, and W. Richardson, Suitability of the Normal Density Assumption for Processing Multispectral Scanner Data, IEEE Transactions on Geoscience Electronics, Vol. GE-10, No. 4, 1972, pp. 158-165.
4. R. B. Crane and W. Richardson, Performance Evaluation of Multispectral Scanner Classification Methods, Proceedings of the Eighth International Symposium on Remote Sensing of Environment, Vol. II, Willow Run Laboratories, Ann Arbor, October 1972, pp. 815-831.
5. R. F. Nalepka, Investigation of Multispectral Discrimination Techniques, Technical Report 2264-12-F, Willow Run Laboratories, Ann Arbor, 1970, pp. 149-172.

Page intentionally left blank



NASA/Langley Research Mail Stop 470 Hampton, Virginia 23365 ATTN: Mr. William Howle	(1)	Department of Watershed Sciences Colorado State University Fort Collins, Colorado 80521 ATTN: Dr. James A. Smith	(1)
U.S. Geological Survey Branch of Regional Geophysics Denver Federal Center, Building 25 Denver, Colorado 80225 ATTN: Mr. Kenneth Watson	(1)	Lockheed Electronics Co. 16811 El Camino Real Houston, Texas 77058 ATTN: Mr. R. Tokerud	(1)
NAVOCEANO, Code 7001 Naval Research Laboratory Washington, D.C. 20390 ATTN: Mr. J. W. Sherman, III	(1)	TRW System Group Space Park Drive Houston, Texas 77058 ATTN: Dr. David Detchmندی	(1)
U.S. Department of Agriculture Administrator Agricultural Stabilization and Conservation Service Washington, D.C. ATTN: Mr. Kenneth Frick	(1)	IBM Corporation 1322 Space Park Drive Houston, Texas 77058 ATTN: Dr. D. Ingram	(1)
Pacific Southwest Forest & Range Experiment Station U.S. Forest Service P.O. Box 245 Berkeley, California 94701 ATTN: Mr. R. C. Heller	(1)	S&D—DIR Marshall Space Flight Center Huntsville, Alabama 35812 ATTN: Mr. Cecil Messer	(1)
Pacific Southwest Forest & Range Experiment Station U.S. Forest Service P.O. Box 245 Berkeley, California 94701 ATTN: Dr. P. Weber	(1)	Code 168-427 Jet Propulsion Laboratory 4800 Oak Grove Drive Pasadena, California 91103 ATTN: Mr. Fred Billingsley	(1)
NASA/Johnson Space Center Mission Planning & Analysis Division Houston, Texas 77058 ATTN: Mr. H. G. De Vezin/FM8	(1)	NASA/Johnson Space Center Technical Library Branch Houston, Texas 77058 ATTN: Ms. Retha Shirkey/JM6	(4)
University of Texas at Dallas Box 688 Richardson, Texas 75080 ATTN: Dr. Patrick L. Odell	(1)	NASA Headquarters Washington, D.C. 20546 ATTN: Mr. W. Stoney/ER ATTN: Mr. Leonard Jaffe/ER ATTN: Mr. M. Molloy/ERR ATTN: Mr. G. Thorley/ERR	(1) (1) (1) (1)
Department of Mathematics University of Houston Houston, Texas 77004 ATTN: Dr. Henry Decell	(1)	Ames Research Center National Aeronautics and Space Administration Moffett Field, California 94035 ATTN: Dr. I. Poppoff	(1)
Institute for Computer Services and Applications Rice University Houston, Texas 77001 ATTN: Dr. M. Stuart Lynn	(1)	Goddard Space Flight Center National Aeronautics and Space Administration Greenbelt, Maryland 20771 ATTN: Mr. W. Nordberg, 620 ATTN: Mr. W. Alford, 563	(1) (1)
U.S. National Park Service Western Regional Office 450 Golden Gate Avenue San Francisco, California 94102 ATTN: Mr. M. Kolipinski	(1)	Lewis Research Center National Aeronautics and Space Administration 21000 Brookpark Road Cleveland, Ohio 44135 ATTN: Dr. Herman Mark	(1)
U.S. Department of Agriculture Statistical Reporting Service Washington, D.C. 20250 ATTN: D. H. VonSteen/R. Allen	(2)	John F. Kennedy Space Center National Aeronautics and Space Administration Kennedy Space Center, Florida 32899 ATTN: Mr. S. Claybourne/FP	(1)
U.S. Department of Agriculture Statistical Reporting Service Washington, D.C. 20250 ATTN: Mr. H. L. Trelogan, Administrator	(1)	NASA/Langley Mail Stop 214 Hampton, Virginia 23665 ATTN: Mr. James L. Raper	(1)